

# SED FITTING WITH MARKOV CHAIN MONTE CARLO: METHODOLOGY AND APPLICATION TO $Z=3.1$ $\text{Ly}\alpha$ -EMITTING GALAXIES

VIVIANA ACQUAVIVA<sup>1</sup>, ERIC GAWISER<sup>1</sup>, LUCIA GUAITA<sup>2,3</sup>

## ABSTRACT

We present GalMC, a MCMC algorithm designed to fit the spectral energy distributions (SED) of galaxies to infer physical properties such as age, stellar mass, dust reddening, metallicity, redshift, and star formation rate. We describe the features of the code and the extensive tests conducted to ensure that our procedure leads to unbiased parameter estimation and accurate evaluation of uncertainties. We compare its performance to grid-based algorithms, showing that the efficiency in CPU time is  $\sim 100$  times better for MCMC for a three dimensional parameter space and increasing with the number of dimensions. We use GalMC to fit the stacked SEDs of two samples of Lyman Alpha Emitters (LAEs) at redshift  $z=3.1$ . Our fit reveals that the typical LAE detected in the IRAC  $3.6\ \mu\text{m}$  band has age =  $0.67\ [0.37 - 1.81]$  Gyr and stellar mass =  $3.2\ [2.5 - 4.2] \times 10^9\ M_{\odot}$ , while the typical LAE not detected at  $3.6\ \mu\text{m}$  has age =  $0.06\ [0.01 - 0.2]$  Gyr and stellar mass =  $2\ [1.1 - 3.4] \times 10^8\ M_{\odot}$ . The SEDs of both stacks are consistent with the absence of dust. The data do not significantly prefer exponential with respect to constant star formation history. The stellar populations of these two samples are consistent with the previous study by Lai et al, with some differences due to the improved modeling of the stellar populations. A constraint on the metallicity of  $z=3.1$  LAEs from broad-band photometry, requiring  $Z < Z_{\odot}$  at 95% confidence, is found here for the first time.

## 1. INTRODUCTION

A galaxy's Spectral Energy Distribution (SED) contains information about its stellar population age, mass, star formation rate (SFR), dust content, and metallicity. Different physical processes leave their imprint in different parts of the spectrum; the wider the wavelength coverage, the more robust the interpretation of the various features of the SED is in terms of galaxy properties. The rest-frame ultraviolet (UV), optical, and infrared (IR) parts of the spectrum offer the combination of depth and angular resolution needed to obtain individual photometry for all galaxies detected in a particular band, which is not generally possible in the far-infrared through sub-millimeter wavelengths where re-emission by dust dominates the luminosity. Spectroscopic information is generally required to determine metallicity and further probes of chemical enrichment, but spectroscopy of unbiased samples of dim galaxies is very difficult to obtain. This leaves photometric UV-through-IR SEDs as the most readily available probe of galaxy properties.

SED fitting is the procedure of comparing models to the observed galaxy SED (for reviews, see Gawiser 2009; Walcher et al. 2010 and references therein). Since the physical properties of the models are known, those of the data can be derived by maximizing the resemblance between data and models. The success and reliability of this method depend on the quality of the available template spectra and on the robustness of the fitting algorithm. The ingredients of the spectral templates typically are libraries of stellar spectra and sets of evolutionary tracks, allowing one to compute the ini-

tial spectrum of a collection of stars of different mass and to follow its evolution with time. Measuring and calibrating these quantities is not an easy process, and the available models cannot uniformly cover stellar populations of all ages, masses, or chemical content. The large variance intrinsic in the spectral templates is reflected in the large number of available stellar population synthesis (SPS) codes, including but not limited to PEGASE (Fioc & Rocca-Volmerange 1997), Maraston (Maraston 1998, 2005), GRASIL (Silva et al. 1998), Starburst99 (Leitherer et al. 1999; Vázquez & Leitherer 2005), GALAXEV (Bruzual and Charlot 2003, Charlot and Bruzual 2010), and GALEV (Kotulla et al. 2009). The scatter among the predictions of different SPS models is in itself an important source of systematic uncertainty (see *e.g.*, Kannappan & Gawiser 2007; Conroy & Gunn 2010 and our discussion in Sec. 4).

Often, the comparison between model and data is made through a  $\chi^2$  minimization, which provides us with a best-fit model. If the probability distribution of the parameters is close to a Gaussian (*e.g.*, Larson et al. 2010), the best fit is a good estimate of the expectation value for each parameter, and the corresponding uncertainties can be evaluated from relative likelihoods assuming a Gaussian profile. This approach can, however, be very dangerous in astronomy. In fact, because of the unprecedented volume of available data, we can now hope to explore new aspects of the physics of galaxies, and this exciting perspective comes at the price of not having prior knowledge of the probability distribution we set out to investigate. Furthermore, the high degree of correlation between astrophysical processes makes probability distributions highly non-Gaussian. In a multi-dimensional space with heavy parameter degeneracies, the “best fit” spectrum is an overly aggressive compression of the information available in the SED, and the assumption of a Gaussian

<sup>1</sup> Department of Physics and Astronomy, Rutgers, The State University of New Jersey, Piscataway, NJ 08854

<sup>2</sup> Departamento de Astronomía y Astrofísica, Universidad Católica de Chile, Santiago, Chile

<sup>3</sup> Institutionen för Astronomi, Stockholms Universitet, SE-106 91 Stockholm, Sweden

likelihood is unjustified. As a result, it is necessary to switch to more sophisticated and robust statistical analysis techniques, such as Markov Chain Monte Carlo (MCMC) algorithms. While the use of simple  $\chi^2$  minimization techniques is still very common, MCMC is becoming a popular statistical method within the SED fitting community (*e.g.*, Panter et al. 2003; Sajina et al. 2006; Nilsson et al. 2007, 2010; Conroy & Gunn 2010; Serra et al. 2011; Pirzkal et al. 2011, Johnson et al. 2011 in prep); however, none of these algorithms is publicly available yet. Our MCMC code, GalMC, is available at <http://www.physics.rutgers.edu/~vacquaviva/web/GalMC.html>.

This paper is organized as follows. In Sec. 2 we present the general features of MCMC algorithms and describe in the detail GalMC, our MCMC code for SED fitting. Sec. 3 illustrates the extensive tests we performed on Mock catalogs to ensure that the algorithm works correctly and that our inferred parameter values and uncertainties are trustworthy. In Sec. 4 we present the results we obtain for two samples of Lyman Alpha Emitters (LAE) at redshift  $z = 3.1$ , for various assumptions of star formation histories (SFH) and stellar populations synthesis (SPS) models; we also compare our results to those obtained for the same samples, but with a different technique, by Lai et al. 2008 (hereafter L08). We summarize our findings in Sec. 5.

## 2. THE MCMC ALGORITHM FOR SED FITTING

The process of parameter fitting involves two essential steps. First, we want to know what is the most representative value, or estimate, of the true value of each parameter included in our fit. Second, we want to evaluate the *uncertainties*, *i.e.*, how likely it is that the true value lies in a given interval in the vicinity of that estimate. In order to accomplish this task, one needs to determine the probability density function of the parameters,  $p(\mathbf{x})$ , which can be used to compute the expectation value of any function of interest as

$$\langle f(\mathbf{x}) \rangle = \int d\mathbf{x} f(\mathbf{x}) p(\mathbf{x}). \quad (1)$$

In many cases where the probability density function is complicated and the dimensionality  $n$  of the parameter space is high, finding alternate ways of evaluating the above integral is required. This can be achieved by informed *sampling* of the parameter space. A sample is a vector of coordinates in parameter space; the objective of the sampling process is to obtain a set of  $R$  *independent* samples  $r_i$  whose distribution is identical to that of the probability density function  $p(\mathbf{x})$ . In practice, this can be achieved if the density of sampled points is proportional to  $p(\mathbf{x})$ . Once this collection of samples is obtained, it will by definition satisfy the property

$$\langle f(\mathbf{x}) \rangle = \int d\mathbf{x} f(\mathbf{x}) p(\mathbf{x}) \simeq \frac{1}{R} \sum_{i=1}^R f(r_i) \quad (2)$$

for any function  $f$ , and performing integrals of the form of Eq. (1) becomes a trivial matter. Often, we will be interested in recovering the *posterior* probability distribution of parameters. According to Bayesian statistics, which we adopt, this depends on the product of the likelihood of observing the data, if that model were the truth,

and on our beliefs on the distribution of parameters, or priors.

A flexible method allowing one to create samples with minimal assumptions on the underlying probability distribution of the parameters with respect to the data is to use Monte Carlo Markov Chains (see *e.g.*, MacKay 2003; Neal 1993; Lewis & Bridle 2002a; Verde et al. 2003; Hajian 2007; Hobson et al. 2010). Monte Carlo methods are computational techniques that make use of random numbers. In this context, a random number is involved in the process of moving from one point in parameter space to the next one. This succession of steps is called a chain. Because the transition to the next step only depends on the current point, having no memory of the previous points, the chain is called a *Markov chain*.

How can we obtain the desired independent samples from a Markov chain? The chain needs to be built satisfying certain rules that restrict the *transition probability*  $T(\mathbf{x}, \mathbf{y})$ , the probability of moving from a point  $\mathbf{x}$  to the point  $\mathbf{y}$ . Two things are required (*e.g.*, MacKay 2003):

1. The desired probability distribution  $p$  is an invariant distribution of the chain:

$$p(\mathbf{x}) = \int d\mathbf{y} T(\mathbf{x}, \mathbf{y}) p(\mathbf{y}). \quad (3)$$

2. The chain is *ergodic*, *i.e.*, the probability distribution of the chain tends to an invariant distribution of the chain, no matter what the initial conditions are.

If these conditions are satisfied, after an initial period where the walk of the chain depends on the initial conditions (the *burn-in* phase), the chain will start sampling from the desired probability distribution (the posterior). However, subsequent steps of the chain are strongly correlated with each other. It is necessary to select points which are distant enough from each other (*thin* the chain) in order to achieve the independent samples needed in Eq. 2.

There are two main tasks associated with the creation of chain. The first is how to explore the parameter space, namely, how to go efficiently from one point to the next one and how to decide whether or not the step just taken will become part of the chain. The latter process generally involves evaluating how similar the model corresponding to a point in parameter space is to the data one wants to fit, based on the model's likelihood given the observed data, and on the chosen priors. The second task is to compute the likelihood of a given model at each point visited by the chain; we do so by using stellar population synthesis models to predict the observable quantities as a function of the input parameters. These two processes are described in detail below.

### 2.1. Description and features of sampling algorithm

The structure of our code, *GalMC*, is deeply indebted to the publicly available algorithm CosmoMC (Lewis & Bridle 2002a). We use a Metropolis sampling algorithm (Hastings 1970), where a trial step  $\mathbf{x}_0 + \Delta\mathbf{x}$  from the current position in parameter space  $\mathbf{x}_0$  is accepted with probability

$$P_{\text{acc}} = \min \left( 1, \frac{p(\mathbf{x}_0 + \Delta\mathbf{x})}{p(\mathbf{x}_0)} \right) \quad (4)$$

where  $p(\mathbf{x}_0)$  is the posterior probability evaluated at  $\mathbf{x}_0$ .

The trial step  $\mathbf{x}_0 + \Delta\mathbf{x}$  is a random deviate drawn from a distribution,  $q(\mathbf{x})$ , known as the proposal density. In principle, Metropolis sampling converges for a wide variety of proposal densities; however, in practice, performance is greatly improved upon judicious selection of this distribution. An important indicator of the performance of the sampling algorithm is the *acceptance rate*, the ratio between the number of accepted steps and the number of trial steps taken; the rule of thumb for optimal acceptance rate is  $\simeq 25\%$  (Roberts et al. 1997). If no previous information on the covariance of parameters (for example, from previous MCMC runs) is available, the proposal density is chosen according to the method of Lewis & Bridle (2002b). A random orthonormal basis in parameter space is chosen; the normalization to orthonormal parameters relies on an initial guess on the width of the target posterior distribution in each direction. Starting from the first direction of the basis vectors, a step of distance  $r$  in that direction is drawn from a proposal density consisting of the sum of a Gaussian and an exponential:

$$q(r) = g\left(\frac{r}{s}\right)e^{-(r/s)^2} + (1-g)e^{-r}; \quad (5)$$

Combining an exponential tail to a Gaussian proposal density is good practice when target probability distributions are not expected to be Gaussian. The factor  $g$  identifies the degree of mixing; we chose the value of  $g = 2/3$  suggested in Lewis & Bridle (2002b) and found nearly optimal acceptance rate, as discussed later in the text. The quantity  $s$  regulates the size of a typical trial step with respect to the (estimated) width of the posterior in that direction, and we choose the value  $s = 2.4$  as suggested by Gelman et al. (1996) and Dunkley et al. (2005). We repeat this procedure taking subsequent trial steps in each direction of the basis vectors, then generate a new random basis, and iterate this process.

More efficient sampling is achieved if the covariances between the parameters are known. When a covariance matrix is available, eigenvectors of this matrix are used to select the basis of orthonormal parameters. We introduce the possibility of using an adaptive covariance matrix. This feature is useful in particular for parameter spaces where the shape of the target distribution depends strongly on the position in parameter space, so that previously obtained covariance matrices might not be a good representation of the covariance in the present case. When this option is selected, the sampling process is stopped at intervals that can be chosen by the user, a new covariance matrix is computed on the desired fraction of the chain (for example, the second half of the chain; typically one would not want to use all the samples to exclude the burn-in region), and the new matrix is used as a subsequent input for the proposal density.

One of the most critical issues in MCMC sampling is the assessment of *convergence*, namely determining whether the probability distribution inferred from the chains resembles the true one. Two key requirements are that the entire relevant region of parameter space has been explored, and that the target distribution from the samples has become stable (*i.e.*, it would not change upon further sampling). The first point is relevant in particular in the presence of a multi-modal distribution;

while in general the MCMC sampling should be able to correctly explore such spaces, it is wise to run multiple chains starting from different locations of parameter space, and we adopt this approach. Examples of algorithms that test convergence on single and multiple chains respectively are the Raftery and Lewis (hereafter RL) statistics (Raftery & Lewis 1992) and the Gelman and Rubin “R” test, comparing the variance of the mean within and between chains (Gelman & Rubin 1992). All these tests are routinely performed by the publicly available software GetDist (Lewis & Bridle 2002a), which we use to analyze the chains. Convergence tests are meant to be performed on an unchanging proposal density. When the adaptive covariance matrix option is selected, the user can choose to stop changing the proposal density either when the desired acceptance rate is achieved (the default value is 25%), or after a certain number of consecutive attempts to update the proposal density without a significant improvement of the acceptance rate. GalMC records the numbers of steps taken up to this time, so that they can be discarded in the computation of the posterior probability and in convergence tests.

For the analysis of data on the LAEs presented in this work, we use the information from the RL statistics to discard the burn-in of the chain and to obtain independent samples by thinning the chain by an appropriate factor. To ensure that convergence has been reached, we require the rather stringent constraint from the R test  $R-1 < 0.02$  and follow the other general guidelines given in Dunkley et al. (2005).

## 2.2. From parameters to observables

To compute the likelihood associated to a point in parameter space (in Bayesian terms, the likelihood of the data, given a model), we need to predict how the observed data would change as a function of the galaxies’ physical properties that we want to measure. The parameter space we have so far explored is defined by the age since the onset of star formation, Age, the total stellar mass processed into stars,  $M_{\text{tot}}$ , the dust reddening, defined by the excess color  $E(B-V)$ , the metallicity in units of the Solar one,  $Z/Z_{\odot}$ , the redshift,  $z$ , and the e-folding time  $\tau$  for exponential star formation history models. In the analysis presented in this paper the redshift has been fixed, since it is well determined for narrow-band selected LAEs. Any combination of SED parameters can be explored, with the exception of  $\tau$  and metallicity. Varying each of these parameters requires the use of a library of templates, and combining them requires too high memory usage to be convenient for an ordinary laptop computer. This does not affect our analysis significantly since our data do not show a strong preference for ESF vs CSF. Memory usage optimization for this purpose will be explored in a subsequent paper.

### 2.2.1. Stellar population synthesis models

As a starting point, one can use stellar population synthesis (SPS) models that predict the rest-frame flux as a function of wavelength for different ages, masses, metallicities, and star formation histories. We use the latest version of the publicly available GALAXEV code (Charlot & Bruzual 2010, private communication; hereafter CB10), although GalMC also supports the earlier version of the same code (Bruzual & Charlot 2003a,

hereafter BC03). The main difference between the two versions is due to the different treatment of the TP-AGB evolution of low- and intermediate-mass stars, whose contribution produces redder rest-frame near-IR colors in the BC10 models (S. Charlot 2010, private communication). We modified the source code of CB10 to reduce running time, which is essential for MCMC algorithms since tens of thousands of iterations are typically required. The main changes include a significant reduction of output written to files by the program `csp_galaxev` and the introduction of a new analytic option for exponentially increasing star formation rate (previously available only through the tabulated star formation rate option, which is considerably slower).

### 2.2.2. Star formation history

Three classes of star formation histories (SFH) can be specified: simple stellar population (SSP), corresponding to an instantaneous burst; constant star formation rate (CSF), and exponential star formation rate (ESF)  $\psi(t) = 1/|\tau|e^{-t/\tau}$ . In the latter case, the e-folding time is also a parameter of the SED fitting, and is allowed to be either positive or negative to represent exponentially decreasing or increasing star formation histories. For the galaxy samples considered in this paper, the allowed choices of star formation histories allow us to find good fits to the data, but in general it would be ideal to measure the star formation history of a galaxy, rather than assuming a functional form for it. This method is employed, *e.g.*, by the algorithms MOPED (Heavens et al. 2000) and VESPA (Tojeiro et al. 2007), where the star formation rate is recovered from the data in several age bins, and the data are also used to constrain the number of underlying stellar populations. We plan to explore this possibility in future application of GalMC, since we expect that high S/N data are needed in order to obtain meaningful constraints on the SFH.

We minimized the number of calls to `csp_galaxev`, which is computationally expensive, by introducing an “adaptive library” for exponential star formation history models. The user can specify the location of a library where the model files are stored; at each step the value of  $\tau$  is compared to those of the models available in the library, and `csp_galaxev` is run only if no close enough model is present. For the analysis in this paper, we have required an accuracy in  $\tau$  of  $\min(2\%, 2 \text{ Myr})$ , but the criterion can be selected by the user. Each time a new model is computed, it is automatically stored in the library. The final product of the call to `csp_galaxev` is a file containing the model as a function of age and star formation history; if required, the code also outputs a file containing the evolution of the stellar mass of the galaxy and a file containing the number of Lyman photons as a function of time. These files are used as input by the program `galaxev.pl`, which extracts the spectrum at the relevant age and normalizes it to the chosen mass value. Thanks to these modifications, the average time per iteration of the GALAXEV part of our MCMC code is  $\sim 0.3$  seconds on a 2.66 GHz MacBook Pro, a factor of  $\sim 20$  faster than the publicly available GALAXEV code.

### 2.2.3. Metallicity

The library of models available through GALAXEV comprise seven metallicity values, between  $Z/Z_{\odot} = 0.005$

and  $Z/Z_{\odot} = 5$ . We allow the user to select a fixed value of metallicity different from any of the templates, or to include metallicity as one of the SED fit parameters. To compute the model spectrum for any value of  $Z$ , we interpolate between the two bracketing values available in the CB10 or BC03 stellar libraries. We tested both linear and logarithmic interpolation, finding that logarithmic interpolation (which became our method of choice) is more reliable. Our MCMC SED fitting runs are technically sound; yet, some caveats need to be mentioned. The dependence of an SED on metallicity is by its own nature complicated, since different types of stars contribute in different ways and at different epochs. Therefore, the precision of our SED modeling is necessarily limited by the paucity of empirical templates, and this systematic uncertainty should be folded into any metallicity measurement coming from photometric data relying on the same templates. We tested the magnitude of this effect by interpolating between two template values to find the predicted spectrum corresponding to one of the other template values, and found discrepancies of order 10% – 20%, significant yet not unreasonable.

We also explored the effect of using different priors on the metallicity distribution. Our choice was of a uniform prior in  $\log Z/Z_{\odot}$ , motivated by the observed distribution for Damped Ly- $\alpha$  systems (Prochaska et al. 2003); however we also performed the MCMC analysis using a uniform prior in  $Z/Z_{\odot}$  and found a mild dependence of our results on the prior used for the sample with less signal-to-noise. This is unsurprising since in general the choice of priors has more influence when the data have less constraining power. We further discuss this issue in Sec. 4.3.

### 2.2.4. Impact of nebular emission

If desired, the contribution to the model fluxes from nebular emission (from both continuum and lines) can be included. The strength of both continuum and lines is assumed to be proportional to the rate of H-ionizing photons per second,  $Q_0$ . Following Schaerer & Vacca (1998), we describe the flux from the continuum emission as

$$f_{\lambda} = \frac{c}{\lambda^2} \frac{\gamma(\lambda)}{\alpha_B} e_{\gamma} Q_0 \quad (6)$$

where  $c$  is the speed of light,  $\alpha_B$  is the case B recombination coefficient for hydrogen,  $1 - e_{\gamma}$  is the escape fraction of Lyman photons, and  $\gamma$  is the total continuum emission coefficient. After choosing a nominal electron density  $N_e$  and temperature  $T_e$  of the emitting gas, and a nominal helium to hydrogen ratio, the emission coefficients of the free-free, free-bound and two-photons continuum for hydrogen and helium can be found in Aller (1984) and Ferland (1980). Following Schaerer & de Barros (2009), we use the values  $e_{\gamma} = 1$ ,  $N_e = 100 \text{ cm}^{-3}$ ,  $T_e = 10000 \text{ K}$ , and  $[\text{He}/\text{H}] = -1$  throughout the paper; similar values were also recently measured for a star-forming galaxy at  $z \sim 2$  by Rigby et al. (2011). The corresponding template is added to the reference spectrum. We have checked that the evolution of the number of Lyman photons and the total emission computed in this way are consistent with those output by the publicly available code for computation of ionizing fluxes StarBurst99 (Leitherer et al. 1999; Vázquez & Leitherer 2005). To add the contribution of emission lines, we assume that the luminosity of

the  $H\beta$  line is given by

$$L(H\beta) = 4.76 \times 10^{-13} e_{\gamma} Q_0 \quad (7)$$

and we use empirical relative line intensities for H, He, C, N, O, S as a function of metallicity (D. Schaerer, private communication and Schaerer & de Barros 2009; data from Anders & Fritze-v. Alvensleben 2003 and Storey & Hummer 1995). This simple treatment is unlikely to capture the complex radiative transfer physics of the Lyman- $\alpha$  line. Since the narrow-band technique allows one to compute the real Lyman- $\alpha$  flux (as the line excess with respect to the continuum), L08 subtracted its contribution from the data and therefore we do not include it in our nebular emission templates.

### 2.2.5. Galactic and IGM absorption

The spectra thereby obtained can be corrected for absorption due to dust in the galaxy and to neutral hydrogen in the intergalactic medium (IGM). The former requires the assumption of a dust law to connect the SED fit parameter  $E(B-V)$  to an optical depth curve. GalMC currently supports two options: the Calzetti law (Calzetti et al. 1994), where the parameter  $R_v$ , related to the size and geometry of dust grains, can be chosen by the user, and a Milky-Way type law (Cardelli et al. 1989). The former is used in the analysis of the present paper, with a value  $R_v = 4.05$  (Calzetti et al. 2000). Starlight absorption by the IGM is taken into account using the prescription from Madau (1995).

### 2.2.6. Comparison with data points

To obtain the spectra in the observed frame we compute the luminosity distance for the cosmology specified by the user through the present day matter density relative to the critical ( $\Omega_{m,0}$ ) and the Hubble constant ( $H_0$ ); a flat geometry ( $\Omega_{\text{tot}} = 1$ ) is assumed. We use the luminosity distance to convert model luminosity to flux; the model fluxes are then convolved with the filter transmission curves specified by the user for each photometric band. Finally, the flux densities  $f_{\nu}$  are obtained as the number of photons corresponding to the fluxes of the model divided by the reference number of photons obtained for a flat spectrum of  $f_{\nu} = 1 \mu\text{Jy}$ :

$$f_{\nu}^i = \frac{\int_{\lambda_{\min}}^{\lambda_{\max}} T^i(\lambda) f_{\lambda}(\lambda) \frac{c}{\lambda} d\lambda}{\int_{\lambda_{\min}}^{\lambda_{\max}} T(\lambda) \frac{c}{\lambda} d\lambda} \quad (8)$$

where  $T^i(\lambda)$  is the transmission curve in the  $i$ -th band between  $\lambda_{\min}$  and  $\lambda_{\max}$ . The likelihood  $\mathcal{L}$  for each model can now be computed (up to a normalization factor, which is irrelevant because we are interested in ratios of likelihoods) as a function of the  $\chi^2$ :

$$-\log \mathcal{L} \propto \chi^2 = \sum_i \frac{(f_{\nu}^i - \phi^i)^2}{\sigma_i^2} \quad (9)$$

where  $\phi^i$  is the flux in the  $i$ -th data point and  $\sigma_i$  is its photometric error.

## 3. ALGORITHM TESTING

In this section we present the extensive tests we performed to ensure that our MCMC SED fit procedure produces reliable results. This goal can be achieved by gen-

erating mock galaxy catalogs with known physical properties and checking that the input parameters are correctly recovered by means of SED fitting. It is necessary to show not only that the estimates of the expectation values of parameters are unbiased, but also that the associated uncertainties, often quoted in terms of 68% and 95% confidence levels, are correct.

Mock catalogs are built in the following manner. We use the modified BC10 SPS code to generate the spectrum of a mock galaxy in 13 bands from observed-frame UV to observed-frame IR and follow the procedure described in Sec. 2.2 to obtain the corresponding flux density in each band. To mimic a photometric uncertainty of 5% in each band, we add a random Gaussian noise of this  $1\sigma$  amplitude to the fluxes. While these assumptions on the uncertainty in the flux and the extent of the available photometry are realistic for our stacked fluxes of LAEs (Gawiser et al. 2007; Lai et al. 2008; Guaita et al. 2010), the results of this test do not depend on the particular numbers.

We build and test mock catalogs for both constant and exponential star formation histories. For the constant star formation case, the running time of the MCMC SED fitting code is limited; a chain of composed by a few thousands steps can be obtained in approximately two hours on a 2.66 GHz MacBook Pro laptop computer, and we have found that in this simple case, when three parameters describe the SED, this is enough to achieve convergence. This makes the CSF scenario a suitable test case to check our marginalized probability distribution of parameters, since this task requires running GalMC on a large number of Mock catalogs, as explained below. Mock catalogs with exponential star formation history are used as a means to test robustness to the presence of degeneracies in the SED and the effect of using different variables and priors in describing the star formation history. Depending on the sign and ratio of age and  $\tau$ , the timescale associated to the exponential rise or decline of the SFR in the mock model, the posterior probability of parameters can be multi-modal or even flat if true degeneracies are present. By testing the MCMC SED fit in such scenarios, we ensured that the input parameters are correctly recovered even in the presence of degeneracies and that we don't overestimate the degree of belief in our results (for example, by assuming that chains have converged when this is not the case). Furthermore, we explored different parameterizations in  $\tau$  and were able to choose tFphe one that leads to the most reliable results.

### 3.1. Constant Star Formation Models: Parameter and error estimation recovery

Our reference model is a galaxy characterized by constant star formation history, Age = 180 Myr, total mass converted into stars  $M_{\text{tot}} = 2.95 \times 10^8 M_{\odot}$ , and excess color  $E(B-V) = 0.2$ . We build 200 mock realizations of this galaxy, adding a Gaussian random scatter to each data point of  $1\sigma$  amplitude equal to the 5% photometric error. We run the MCMC SED fitting code on the reference model (without scatter in the photometry), ensuring that we correctly recover the input parameters, and hereby obtaining a reliable covariance matrix which gives a nearly optimal acceptance rate (around 30%). We then run the MCMC SED fitting code on each of

the mock catalogs, using this covariance matrix as an input for the proposal density. After discarding the chains that have not yet converged after 10000 steps, we are left with 172 reliable runs that we use to test the recovery of uncertainties. To do so, we compute the frequency with which each of the “true” (input) parameters is within the 68% and 95% confidence levels from the 1-D marginalized posterior probability distributions obtained for the mock catalogs. The  $i$ -dimensional *marginalized* posterior distribution of parameters is obtained by integrating the  $n$ -dimensional posterior distribution  $p$  in  $n-i$  dimensions; for example, for CSF the 1-D age marginalized posterior probability is given by the function:

$$p(\text{Age}) = \int dM_{\text{tot}} \int dE p(\text{Age}, M_{\text{tot}}, E). \quad (10)$$

Since a thinned MCMC chain has the property that the density of points is proportional to the posterior probability (the product of likelihood and priors), the above integral is easily computed as a sum over the points of the chain that takes into account the time spent at each location in parameter space. By definition, we expect that the true values are within the region allowed at 68 (95)% confidence 68 (95)% of the time. We find that this is the case for each parameter, within the Poisson fluctuation error (order of 8% effect) associated with our statistics. We conclude that the uncertainties we report are reliable. We note that since we use the well-tested GetDist software to compute the posterior probability distributions and to perform convergence tests, this check ensures that our sampling algorithm and calculation of likelihood are implemented correctly.

A more immediate visualization of the match between the observational errors and the corresponding uncertainties in the parameters can be obtained by plotting the probability distribution of the parameters in the “true” model (without the scatter), and the scatter of the best-fit values obtained for the 172 mock catalogs. The agreement between these two quantities is however only expected in the case of perfect (one-to-one) correspondence between the likelihood associated to the data for each model and the posterior probability, and assuming that the best fit represents the truth for each model. These assumptions are not unreasonable for these simple models, leading to the observed agreement shown in Fig. 1.

### 3.2. Exponential Star Formation Models: Parameter recovery

Models with exponentially decreasing star formation rate have been often used in the literature, and recently increasing exponential star formation rates have also been considered (Maraston et al. 2010; Guaita et al. 2010; Lee et al. 2010). While studying the constraints on these models is interesting, they are often affected by degeneracies intrinsic in the model SEDs. For this reason, it is necessary to pay particular attention to the use of suitable sampling variables and priors. We have implemented three possible choices of parametrization in  $\tau$ , using as sampling variable  $\tau$ ,  $1/|\tau| \ln(|\tau|)$ , and  $1/\tau$ ; all of these are currently available options in the code. We ran GalMC for each of these choices on a range of Mock catalogs, using different signs and numerical values for the ratio of age and  $\tau$ . The parametrization  $1/\tau$  was

our final variable of choice, since we did not observe any bias or erroneous convergence problem for this case. This description has the attractive feature that the constant star formation rate case, which is the limit for  $\tau \rightarrow \infty$  for both positive and negative values of  $\tau$ , occupies a single spot ( $1/\tau = 0$ ) in parameter space (unlike the other parameterizations). This makes the interpretation of constraints easier and transparent. We implemented it with a flat prior in  $\ln \tau$ . In fact, during the parametrization selection process, we found that the use of a flat prior in  $1/\tau$  led in some cases to a poor recovery of the input parameters, as did the use of a flat prior in  $\tau$  when sampling using  $\tau$  as one of the parameters, which is not uncommon in the literature. On the other hand, no bias or erroneous results were found by us when using a uniform prior in  $\ln |\tau|$ , as shown in Fig. 2, confirming that this is the best choice of prior.

The sensitivity of the results to the choice of priors emphasizes the need for caution when computing constraints on these models. We expect that significant improvement could be achieved by using variables which are closer to the eigenvectors of the covariance matrix; we defer this treatment to a subsequent paper (Acquaviva et al 2011 in prep).

Results obtained for three illustrative mock catalogs, chosen to have different star formation histories (increasing and decreasing),  $\tau/\text{Age}$  ratios, and redshifts, are shown in Fig. 2. The 1-D posterior distributions are obtained by integrating the posterior probability over the remaining  $N-1$  dimensions. In all the Figures, the normalization of these plots is arbitrary; we show it using the usual CosmoMC convention, where the height of the peak is one. We find good agreement, in each case within the region allowed at 68% confidence, between the input parameters and those recovered by means of SED fitting.<sup>4</sup>

### 3.3. Comparison with grid-based techniques

As explained in Sec. 2, the MCMC technique allows one to draw samples from the posterior distribution, and to use them to compute any meaningful quantity associated with it, *e.g.*, the expectation values of parameters, and their uncertainties (for any quantile). This is obviously not the only way to achieve this goal. Why should one make use of the MCMC technique rather than sample the posterior distribution by means, for example, of a fine grid in the parameter space? There are two main arguments in favor of this choice: *efficiency* and *reliability*. The first can be understood as follows. In general, the allowed range of values for each parameter is much larger than the interesting range for that parameter, namely the region where the likelihood is significantly different from zero, which is the relevant one in computing integrals of the form given by Eq. (1). For example, to measure the age of a galaxy, one would typically want to test values from the youngest age that be can resolved (possibly a few million years) to the age of the Universe at the redshift of the galaxy. Even using a logarithmic spacing, this spans several e-foldings. On the other hand, the width

<sup>4</sup> For Mock catalogs 1 and 3 convergence is slow and the stringent criterion  $R - 1 < 0.02$  would not be satisfied yet. This is encouraging since it indicates that our criteria are indeed fairly conservative.

of the interesting region (say, to within 99% confidence level from the best-fit or mean age) will typically be much smaller (of the order of three e-foldings for the example of Fig. 1). By sampling on a grid, a large fraction of time will be spent sampling regions of no interest, while in MCMC, after a burn-in period, every step, whether or not is accepted, is taken in the informative region. For an equal amount of CPU time, this means that the finesse of the results achieved by MCMC - the accuracy when computing *e.g.*, the integral in Eq. (1) - is much higher. As an example, we considered the Mock catalog described in 3.1. We found that in 86% of our 200 test runs, 10000 steps were enough to achieve convergence. Let us assume, to be conservative, that 20000 steps is a safe number of steps for a hypothetical MCMC chain to converge. For this three-dimensional parameter space, this total number corresponds to computing the model SED for  $\sim 27$  values for each parameter. We performed the grid sampling assuming an allowed range which is comparable for  $\log(\text{Age})$  and  $E(B-V)$  to the top-hat priors used for the MCMC case, and significantly narrower in  $\log(\text{Mass})$ . The number of points falling in the relevant region of the posterior (shown in Fig. 1) is only a handful (three or four) for  $\log(\text{Mass})$  and  $E(B-V)$ , and about ten for  $\log(\text{Age})$ . Therefore, the fraction of time spent in the informative region is only  $(4 \times 4 \times 10)/27^3$ , which is less than 1%. This is too coarse a grid to reconstruct the marginalized probability distribution, and an attempt at it via the integration of the posterior leads to a delta function at the best fit found in the grid ( $\chi^2 = 20$ ,  $\text{Age} = 108$  Myr,  $\text{Mass} = 2.62 \times 10^8 M_\odot$ , and  $E(B-V) = 0.23$ , to be compared with the input values  $\text{Age} = 180$  Myr,  $\text{Mass} = 2.95 \times 10^8 M_\odot$ , and  $E(B-V) = 0.2$ ). The situation only worsens for spaces with higher dimensionality, since the number of MCMC steps typically scales linearly in the number of parameters  $N$ , while the number of steps in grid-based methods grows exponentially. For six parameters, the expected numbers of MCMC steps would be of the order of  $6 \times 10^4$ , which corresponds to sampling less than 10 values for each parameter on a grid in the same CPU time.

Reliability is an equally important issue. The spacing on a grid has to be chosen arbitrarily (and, as we saw, it is computationally very expensive to choose a fine pace), and there is no guarantee that one would not miss features of the posterior distribution which are narrower than the interval between adjacent values. Conversely, in MCMC one can use an adaptive step size (such as the one implemented by us by means of the adaptive covariance matrix), sampling more finely in high-likelihood regions. Furthermore, the use of convergence tests and the possibility of running multiple chains provide strong indications that all the interesting regions of parameter space have been adequately explored. These attractive features of MCMC technique allow for a true optimization of CPU time.

## 4. RESULTS

### 4.1. LAE samples

As part of the MUSYC survey (Gawiser et al. 2006), Gronwall et al. (2007) discovered a complete sample of 162 Lyman Alpha Emitters (LAE) at  $z=3.1$  in a narrowband survey of the Extended Chandra Deep Field

South (ECDF-S). The available observed-frame broadband photometry for this sample encompasses six UV-optical bands (U, B, V, R, I, z), two near-IR bands (J and K), and the four IR IRAC bands. L08 eliminated from this sample 86 LAEs in regions of the IRAC images suffering significant contamination from nearby neighbors, and created samples of 18 IRAC-Detected and 52 IRAC-Undetected LAEs, with the “detection” at flux density  $\geq 0.3\mu\text{Jy}$  equating to roughly  $3\sigma$  significance. Six additional IRAC detections were classified as probable AGNs or high-dust galaxies and were analyzed separately. Median-stacked SEDs of these samples were formed and fit using BC03 models; for constant star formation rate, best-fit models showed no dust for either sample and stellar masses of  $9 \times 10^9$  [ $3 \times 10^8$ ]  $M_\odot$  and ages of 1600 [160] Myr for the IRAC-Detected [Undetected] sample; the corresponding uncertainties at 68% confidence are reported in Table 3. The best-fit values obtained for exponentially declining SFH were within 20% of the values obtained for CSF. Gawiser et al. (2007) fit a two-population SED model to the stacked IRAC-Undetected SED, using an exponentially declining SFH for each population. Although their best-fit model placed 80% of the stellar mass in an underlying old population and only 20% in a  $\sim 20$  Myr-old starburst population, they were unable to rule out a single-population fit with age  $\sim 150$  Myr, in good agreement with the results of L08. In the following, we investigate further the median-stacked SEDs of the L08 IRAC-Detected and IRAC-Undetected samples of  $z = 3.1$  LAEs; we refer to these two samples as “IRAC Det” and “IRAC Und” in figures and tables. Our treatment differs from the previous analyses in the use of the full probability distribution of parameters obtained from the MCMC SED fitting code, as well as in the inclusion of the contribution of nebular continuum and emission lines and in the use of the improved CB10 SPS models.

### 4.2. Physical properties and SFH

We begin our MCMC analysis considering a four-dimensional parameter space defined by age, dust reddening, metallicity, and stellar mass, and assuming constant star formation history. The dust reddening is parametrized by the excess color  $E(B-V)$  assuming a Calzetti dust absorption law. The input parameter of our modified GALAXEV code is the total mass of gas turned into stars  $M_{\text{tot}}$  (the integral of the instantaneous star formation rate over the age of the galaxy), but we report constraints on the more meaningful stellar mass  $M_*$ , which takes into account the life cycle of stars and the associated mass loss. This effect is typically of order of 10 – 20%. We assume that none of the gas thereby injected into the IGM is reprocessed to stars. We use top-hat priors on  $\log(\text{Age})$ ,  $\log(M_{\text{tot}})$ ,  $\log_{10}(Z/Z_\odot)$  and  $E(B-V)$ ; the allowed ranges are reported in Table 1. The lower bound of  $10^6$  years for the age comes from an educated guess of the applicability of the SPS models we consider (Bruzual & Charlot 2003b). Our reference cosmology assumes total energy density relative to critical  $\Omega_{\text{tot}} = 1$ , matter density  $\Omega_m = 0.258$ , and Hubble constant  $H_0 = 73$  km/sec/Mpc. In our baseline model we assume a Salpeter Initial Mass Function (IMF) as in L08 and we include the contribution of nebular emission as described in Sec. 2.2.4. While in simple stellar popula-

tion (SSP) models the contribution of emission lines and continuum become negligible after  $\sim 20$  and  $\sim 5$  Myr respectively (*e.g.*, Nilsson et al. 2010), when star formation is ongoing Ly- $\alpha$  photons are continuously produced and this effect cannot be neglected. For example, in CSF models the number of Ly- $\alpha$  photons increases for the first  $\simeq 20$  Myr and stays roughly constant thereafter. It is however true that for older stellar populations the relative importance of nebular emission with respect to the stellar continuum decreases with time.

Results for the IRAC Detected and Undetected samples are presented in Fig. 3 and summarized in Table 2. In the table we report the mean value of parameters from the posterior distribution  $p$ , *e.g.*,

$$\langle \text{Age} \rangle = \int d\text{Age} \int dM_{\text{tot}} \int dE \text{ Age } p(\text{Age}, M_{\text{tot}}, E) \quad (11)$$

where the integral is over the domain specified above. This is easily computed by averaging over the R samples  $r_i$  obtained after thinning the Markov chain:

$$\begin{aligned} \langle \text{Age} \rangle &= \int d\text{Age} \int dM_{\text{tot}} \int dE \text{ Age } p(\text{Age}, M_{\text{tot}}, E) \\ &\simeq \frac{1}{R} \sum_i \text{Age}(r_i) \end{aligned} \quad (12)$$

If the posterior is not symmetric (as in this case), the mean values can differ from the best-fit values; we return to this issue in Sec. 4.3. We also report the 68% uncertainties, computed integrating the posterior from each side toward the high-probability region and stopping when the integral under the curve on each side is 16% of the total (32% in the case of E(B-V), which has a one-tail distribution). In Fig. 3 we show 1-D the marginalized posterior probability distribution for each parameter, and the 2-D marginalized constraints, including contours for the 68 and 95% confidence regions, on the different combinations. Degeneracies between parameters appear here as diagonal axes of these ellipse-like curves; the one between age and stellar mass is evident in the bottom left panel.

The different nature of the two samples is easily seen from Fig. 3. While both stellar populations are consistent with having no or very little dust ( $E(B-V) < 0.04$ ), the sample detected in IRAC is considerably older and more massive than the undetected one, with mean ages and stellar masses of  $0.67 [0.37 - 1.18]$  Gyrs and  $3.2[2.5 - 4.2] \times 10^9 M_{\odot}$  versus  $6 [1-20] \times 10^7$  yrs and  $2 [1.1 - 3.4] \times 10^8 M_{\odot}$  respectively. We are also able to set a constraint on the metallicity of these LAEs. For the IRAC detected sample, we find  $Z/Z_{\odot} = 0.036 [0.005-0.07]$ , while for the IRAC undetected sample,  $Z/Z_{\odot} = 0.05 [0.005-0.13]$ . In both cases, Solar metallicity is excluded at more than 95% confidence. We note however that using a different prior (uniform in  $Z/Z_{\odot}$ ) slightly relaxes the allowed parameter range. The metallicity constraint becomes weaker for the IRAC undetected sample, which has lower S/N; in this case Solar metallicity is only excluded at 68% confidence. This interesting result is in alignment with previous claims that  $z=3.1$  LAEs have metallicity lower than Solar (Finkelstein et al. 2010) and might be galaxies in their first star formation episode (*e.g.*, Hu et al. 1998).

We also consider the effect of assuming an exponential star formation history. In this case there is an additional parameter of the SED fit, the e-folding time  $\tau$ . As explained in Sec. 3.2, we use the variable  $1/\tau$  in the MCMC sampling; we apply a flat prior in  $\log |\tau|$  between the values of  $\tau = -4$  Gyr and  $\tau = 4$  Gyr (for  $|\tau| > 2$  Gyr the SED is indistinguishable from that of a Constant Star Formation History, so that the CSF case is included by this parametrization as the limit of  $1/\tau \rightarrow 0.25 \text{ Gyr}^{-1}$ ). For this analysis, we fix the metallicity at the same value for both samples, in order to isolate the effect of assuming a more general SFH. We chose the value  $Z/Z_{\odot} = 0.02$ , which is consistent with the range found earlier and for which we have an empirically calibrated stellar template available. The results are shown in Fig. 4; we consider the four most meaningful 2D combinations of parameters. We plot the posterior probability as function of  $1/\tau$  to show the output of the MCMC SED fitting algorithm, but we report the constraints in terms of  $\tau$  for clarity in Table 2.

In both cases the preferred value of  $\tau$  is close to the CSF value. For the Undetected sample, we find  $|\tau| > 0.12$  Gyr at 68% confidence, while for the Detected sample the corresponding constraint is  $|\tau| > 0.67$  Gyr. This can perhaps be interpreted in terms of the different age of these two stellar populations. In fact, the SED is sensitive to the ratio  $\text{Age}/\tau$ ; when this ratio is small, there aren't enough e-foldings to distinguish the effect of the exponential SFH from a Constant one. The change in the probability distribution of the other parameters is also shown in Fig. 4. The main effect of the inclusion of an extra parameter is the inability to rule out very old ages (leftmost panels), so that the sharp difference in the age of the two populations found for CSF is somewhat mitigated. The results for the stellar mass and dust reddening are stable to the change in SFH. This was observed previously by L08 for these  $z=3.1$  samples and was also found to be true for redshift  $z=2.1$  LAEs by Guaita et al. (2010). Overall, both the IRAC Detected and Undetected LAEs are reasonably fit by a Constant Star Formation History model characterized by Age, stellar mass, and dust reddening  $E(B-V)$ , with fixed metallicity. The best fit models for the two cases have a  $\chi^2$  of 13.4 and 7.6 respectively, for twelve data points and three free parameters. In the case of exponential SFH, the preferred  $\tau$  values are very close to the ones corresponding to CSF, but there is a modest improvement in the goodness-of-fit coming from adding one extra parameter, with best-fit  $\chi^2$  of 10.8 and 5.2 respectively. These values are obtained for an exponentially increasing SFH. However, CSF is well within the range of e-folding times allowed at 68% confidence for both samples, and we conclude that the inferred physical properties of our LAEs are robust to different assumptions on the SFH and that the CSF model is a satisfying parametrization for both samples.

#### 4.3. Impact of SPS modeling

Besides the choice of star formation history, many features of the stellar population synthesis modeling influence the results of SED fitting. Our preferred model includes nebular emission as described in Sec. 2.2.4, considers metallicity  $Z$  as one of the SED fit parameters, and employs the latest version of the GALAXEV code



and templates, CB10. Here we change these assumptions one by one in a sequence and show their effect on the probability distribution of SED parameters and therefore on the inferred properties of LAEs. We first fix the metallicity at the Solar value, then neglect the effect of nebular emission, and then use the earlier version of the GALAXEV code and templates, BC03. The final combination of assumptions coincides with those used in the analysis of L08.

Our results are shown in Fig. 6 and summarized in Table 3; the last column also shows the best-fit  $\chi^2$ , for twelve data points and four (for varying  $Z$ ) or three free parameters. In all cases the inferred amount of dust reddening is nearly insensitive to the assumptions on the SPS modeling, except for a moderate increase of the uncertainties when metallicity is varied, and we do not discuss it further. This behavior is expected because we subtract the Ly- $\alpha$  flux from our photometry before fitting. As a result, the impact of nebular emission lines and continuum, as well as the difference among the stellar templates we used, is negligible in the rest-UV region of the spectrum (observed frame UV-optical), whose slope measures dust reddening.

The effect of assuming fixed Solar metallicity, as in L08, produces an appreciable effect on the probability distribution of SED parameters for the IRAC Undetected sample. This same trend with metallicity was also observed by L08 for two discrete values of  $Z/Z_\odot$ ; however, a direct comparison is not possible because we use the CB10 models where they used the BC03 ones. The improvement in the best-fit  $\chi^2$  for models of low metallicity is especially marked for the IRAC Detected sample, as shown in Table 3.

Adding the contribution of nebular emission produces little impact on the SED fit of the IRAC Detected sample. In fact, this population is older, and the relative strength of nebular emission lines and nebular continuum with respect to the stellar continuum is lower. Furthermore, the strongest emission lines (H $\alpha$ , H $\beta$  and OIII) mainly affect the K band; the corresponding data point was higher than the continuum SED, so that the SED parameters do not need to change to accommodate this feature, and instead the best-fit  $\chi^2$  improves as a result. This behavior can be seen in the right-hand panel of Fig. 5. On the other hand, the inclusion of nebular emission has a strong effect both on the model SED and on the allowed parameter ranges for the younger LAEs in the IRAC Undetected sample, as seen in the left-hand panel of Fig. 5 and in Table 3. A new peak in the posterior probability distribution, at very young ages of  $\sim 1$  Myr and very low masses around  $3 \times 10^7 M_\odot$ , is present. As a result, the stellar population is interpreted as significantly younger and less massive when nebular emission is taken into account. This response of high-redshift galaxies to SED fitting including nebular emission is in agreement with the findings of Schaerer & de Barros (2009).

Finally, we consider the effect of using the BC03 stellar templates rather than the newer BC10 ones. Because of the differences in the templates discussed in Sec. 2.2.1, when BC03 models are used, older ages and higher stellar masses are needed to fit the SEDs of both samples.

The last set of assumptions in SPS modeling (no nebular emission, solar metallicity, BC03 templates) coincides with the previous analysis of the same samples presented

in L08. This allows a direct comparison of the best-fit approach vs the use of mean values computed from the posterior distribution, shown in Fig. 6 and in Table 3. To compare the values of stellar masses on the same basis, we need to account for the mass loss due to the life cycle of stars, which we consider and was not explicitly quoted in that paper. We do so by multiplying the instantaneous star formation rate by the age reported in L08 to obtain the total mass transformed into stars  $M_{\text{tot}}$ , and transform it into stellar mass using the conversion factor output by our code for the same model. Our best-fit parameters (also shown in the figure as magenta dot-dashed vertical lines) are in exact agreement with those of L08; however, the best-fit values do not lie exactly at the mean of the marginalized probability distribution, leading to a moderate shift of ages and stellar masses. The disagreement is within the 68% confidence level and is expected when the N-dimensional posterior distribution is asymmetric.

## 5. CONCLUSIONS

We built a Markov Chain Monte Carlo code for SED fitting, GalMC, designed to determine physical properties of galaxies including age, stellar mass, dust content, metallicity, star formation history and redshift (the latter was fixed in the present analysis). The purpose of MCMC codes is to recover the probability distribution function (PDF) of the fit parameters by building a number of independent *samples* with the same statistical properties as the PDF. In the Bayesian approach to statistical inference the relevant probability distribution is the *posterior* probability, which depends on the likelihood of the parameters given the underlying data and on our prior beliefs about the model parameters. Samples are built by exploring the parameter space in such a way that the density of sampled points is proportional to the probability distribution that we want to map. Once the samples have been obtained, they can be used to easily calculate the expectation values (*i.e.*, means) of parameters and the associated uncertainties at any confidence level, since integrals in any dimensions become tractable sums. Because most of the CPU time is spent in high-likelihood, informative regions of parameter space, MCMC algorithms offer a substantial improvement in efficiency with respect to methods based on mapping the posterior probability on a grid of reference parameter values. We found that this speed-up factor is of order  $\sim 100$  for a three-dimensional parameter space, and would rapidly increase for a larger number of parameters, since the expected scaling is linear in the number of parameters for MCMC, and exponential for grid-based methods. MCMC codes also offer warning flags of unreliable results by means of convergence tests and allow one to choose the sampling step size adaptively in each direction.

We conducted an extensive series of tests on Mock catalogs built from a variety of stellar populations to ensure that the input parameters and their uncertainties were correctly recovered by GalMC, and to investigate the effect of using different priors. We then used the code to determine the physical properties of two samples of LAEs at redshift  $z = 3.1$ . Our analysis showed that, on average, the LAEs in the IRAC detected sample are older and more massive than their counterparts in the IRAC undetected sample, and that both populations are

essentially dust-free. Furthermore, LAEs at  $z = 3.1$  have metallicities significantly lower than Solar, in agreement with recent spectroscopic studies of high-redshift LAEs. We investigated different assumptions on the star formation history of the LAEs, performing the fit with either constant or exponential (including increasing and decreasing) star formation rate, and found that the constant star formation model is favored by the data. Our results essentially confirm the findings of L08 about the distinct nature of the IRAC Detected and Undetected populations. However, important differences are found with respect to the previous analysis. The use of the improved SPS models from Charlot and Bruzual 2010 (as opposed to those from 2003) shifts the age of the older IRAC-detected population toward younger values, and therefore the mass toward lower values. Further changes are caused by our use of varying metallicity, by the inclusion of the flux from nebular continuum and emission lines, and by the use of the expectation values of parameters computed from the posterior distribution rather than best-fit values. The quoted uncertainties also differ, since we use the Bayesian approach and compute them as the 68% quantile of the marginalized posterior distribution; for the  $i$ -th parameter, this means that we integrate the posterior distribution in all but the  $i$ -th direction.

The algorithm development and SED analysis conducted in the present work set the foundation for a range of future applications. Our analysis showed that not only the parameter expectation values, but also their uncertainties, depend on the assumptions made in the SPS modeling; for example, the inclusion of nebular emission noticeably worsens the ability to rule out young ages for the IRAC Undetected sample. In the present work we have compared two models with differing number of parameters (a constant and an exponential star formation one). For the LAEs samples we studied, the MCMC sampling was directed toward very large values of the parameter  $\tau$ , similar to the CSF value, even when ESF was used, and the improvement in the quality of the fit due to use of one additional parameter was modest. We could conclude that the data do not strongly prefer ESF to CSF models. However, to better quantify this statement, or in general to assess whether the data favor the inclusion of additional parameters, we plan to use model selection (Jeffreys 1961).

SED fitting is a technique of ever-increasing importance in astronomy, and it is essential that the tools used for statistical analysis keep up with the fast pace of the improvement in the data. Large-volume photometric surveys allow us to explore new and exciting directions, and this must be done while avoiding biasing assumptions on the shape of the probability function, and while maximizing the accuracy in the reported uncertainties. Markov Chain Monte Carlo algorithms enable efficient, reliable estimation of parameter expectation values and uncertainties, are suitable for exploring parameter spaces of high dimensionality, and are able to reveal degeneracies among parameters. GalMC is our implementation of this approach, and we hope it will prove useful for a wide range of astrophysical applications.

codes and shell scripts, Antony Lewis and Sarah Bridle for making CosmoMC publicly available and for helpful advice, Gustavo Bruzual and Stephane Charlot for providing and commenting the newest GALAXEV code and templates, Daniel Schaerer for help on the inclusion of nebular emission, Kamson Lai for useful conversations on the photometry of the LAEs, Mike Berry, Nick Bond, Robin Ciardullo, Niv Drory, Harold Francke, Caryl Gronwall, Peter Kurczynski, Kim Nilsson, Nelson Padilla, Kevin Schawinski, Licia Verde, Jean Walker, and Grant Wilson for helpful comments during the completion of this project and on the manuscript draft. We thank the anonymous referee for many expert comments and suggestions that helped us to improve the paper significantly. This material is based on work supported by NASA through an award issued by JPL/Caltech and by the National Science Foundation under grant AST-0807570. Partial financial support for V.A. was provided by the National Science Foundation under the PIRE program (award number OISE-0530095). E.G. thanks the Department of Physics at U.C. Davis for hospitality during the completion of this research.

It is a pleasure to thank Amir Hajian, Ross Fadely and Chuck Keeton for many useful conversations on MCMC

## REFERENCES

- Aller, L. H., ed. 1984, *Astrophysics and Space Science Library*, Vol. 112, Physics of thermal gaseous nebulae
- Anders, P. & Fritze-v. Alvensleben, U. 2003, *A&A*, 401, 1063
- Bruzual, G. & Charlot, S. 2003a, *MNRAS*, 344, 1000
- . 2003b, *MNRAS*, 344, 1000
- Calzetti, D., Armus, L., Bohlin, R. C., Kinney, A. L., Koornneef, J., & Storch-Bergmann, T. 2000, *ApJ*, 533, 682
- Calzetti, D., Kinney, A. L., & Storch-Bergmann, T. 1994, *ApJ*, 429, 582
- Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, *Astrophys. J.*, 345, 245
- Charlot, S. & Bruzual, G. 2010
- Conroy, C. & Gunn, J. E. 2010, *ApJ*, 712, 833
- Dunkley, J., Bucher, M., Ferreira, P. G., Moodley, K., & Skordis, C. 2005, *MNRAS*, 356, 925
- Ferland, G. J. 1980, *PASP*, 92, 596
- Finkelstein, S. L., Hill, G. J., Gebhardt, K., Adams, J., Blanc, G. A., Papovich, C., Ciardullo, R., Drory, N., Gawiser, E., Gronwall, C., Schneider, D. P., & Tran, K. 2010, *ArXiv e-prints*
- Fioc, M. & Rocca-Volmerange, B. 1997, *A&A*, 326, 950
- Gawiser, E. 2009, *New Astronomy Review*, 53, 50
- Gawiser, E., Francke, H., Lai, K., Schawinski, K., Gronwall, C., Ciardullo, R., Quadri, R., Orsi, A., Barrientos, L. F., Blanc, G. A., Fazio, G., Feldmeier, J. J., Huang, J.-s., Infante, L., Lira, P., Padilla, N., Taylor, E. N., Treister, E., Urry, C. M., van Dokkum, P. G., & Virani, S. N. 2007, *ApJ*, 671, 278
- Gawiser, E., van Dokkum, P. G., Herrera, D., Maza, J., Castander, F. J., Infante, L., Lira, P., Quadri, R., Toner, R., Treister, E., Urry, C. M., Altmann, M., Assef, R., Christlein, D., Coppi, P. S., Durán, M. F., Franx, M., Galaz, G., Huerta, L., Liu, C., López, S., Méndez, R., Moore, D. C., Rubio, M., Ruiz, M. T., Toft, S., & Yi, S. K. 2006, *ApJS*, 162, 1
- Gelman, A., Roberts, G., & Gilks, W. 1996, in *Bayesian Statistics*, ed. J. M. Bernardo et al., Vol. 5 (OUP), 599
- Gelman, A. & Rubin, D. B. 1992, *Statistical Science*
- Gronwall, C., Ciardullo, R., Hickey, T., Gawiser, E., Feldmeier, J. J., van Dokkum, P. G., Urry, C. M., Herrera, D., Lehmer, B. D., Infante, L., Orsi, A., Marchesini, D., Blanc, G. A., Francke, H., Lira, P., & Treister, E. 2007, *ApJ*, 667, 79
- Guaity, L., Acquaviva, V., Padilla, N., Gawiser, E., Bond, N., Ciardullo, R., Treister, E., Kurczynski, P., Gronwall, C., Lira, P., & Schawinski, K. 2011, *ApJ*, in press
- Guaity, L., Gawiser, E., Padilla, N., Francke, H., Bond, N. A., Gronwall, C., Ciardullo, R., Feldmeier, J. J., Sinawa, S., Blanc, G. A., & Virani, S. 2010, *ApJ*, 714, 255
- Hajian, A. 2007, *Phys. Rev.*, D75, 083525
- Hastings, W. K. 1970, *Biometrika*, 57, 97
- Heavens, A. F., Jimenez, R., & Lahav, O. 2000, *MNRAS*, 317, 965
- Hobson, M. P., Jaffe, A. H., Liddle, A. R., Mukeherjee, P., & Parkinson, D. 2010, *Bayesian Methods in Cosmology*, ed. Hobson, M. P., Jaffe, A. H., Liddle, A. R., Mukeherjee, P., & Parkinson, D.
- Hu, E. M., Cowie, L. L., & McMahon, R. G. 1998, *ApJ*, 502, L99+
- Jeffreys, H. 1961, *Theory of probability*, 3rd edn. (Oxford University Press)
- Kannappan, S. J. & Gawiser, E. 2007, *ApJ*, 657, L5
- Kotulla, R., Fritze, U., Weibacher, P., & Anders, P. 2009, *MNRAS*, 396, 462
- Lai, K., Huang, J., Fazio, G., Gawiser, E., Ciardullo, R., Damen, M., Franx, M., Gronwall, C., Labbe, I., Magdis, G., & van Dokkum, P. 2008, *ApJ*, 674, 70
- Larson, D., Dunkley, J., Hinshaw, G., Komatsu, E., Nolte, M. R., Bennett, C. L., Gold, B., Halpern, M., Hill, R. S., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Odegard, N., Page, L., Smith, K. M., Spergel, D. N., Tucker, G. S., Weiland, J. L., Wollack, E., & Wright, E. L. 2010, *ArXiv e-prints*
- Lee, S., Ferguson, H. C., Somerville, R. S., Wiklund, T., & Giallisco, M. 2010, *ApJ*, 725, 1644
- Leitherer, C., Schaerer, D., Goldader, J. D., González Delgado, R. M., Robert, C., Kune, D. F., de Mello, D. F., Devost, D., & Heckman, T. M. 1999, *ApJS*, 123, 3
- Lewis, A. & Bridle, S. 2002a, *Phys. Rev.*, D66, 103511
- . 2002b, <http://cosmologist.info/notes/CosmoMC.pdf>
- MacKay, D. 2003, *Information Theory, Inference, and Learning Algorithms* (Cambridge Univ. Press)
- Madau, P. 1995, *ApJ*, 441, 18
- Maraston, C. 1998, *MNRAS*, 300, 872
- . 2005, *MNRAS*, 362, 799
- Maraston, C., Pforr, J., Renzini, A., Daddi, E., Dickinson, M., Cimatti, A., & Tonini, C. 2010, *MNRAS*, 407, 830
- Neal, R. 1993, *Technical Report CRG-TR-93-1*, University of Toronto
- Nilsson, K. K., Möller, P., Möller, O., Fynbo, J. P. U., Michałowski, M. J., Watson, D., Ledoux, C., Rosati, P., Pedersen, K., & Grove, L. F. 2007, *A&A*, 471, 71
- Nilsson, K. K., Östlin, G., Möller, P., Möller-Nilsson, O., Tapken, C., Freudling, W., & Fynbo, J. P. U. 2010, *ArXiv e-prints*
- Panther, B., Heavens, A. F., & Jimenez, R. 2003, *MNRAS*, 343, 1145
- Pirzkal, N., Rothberg, B., Nilsson, K. K., Finkelstein, S., Koekemoer, A., Malhotra, S., & Rhoads, J. 2011, *ArXiv e-prints*
- Prochaska, J. X., Gawiser, E., Wolfe, A. M., Castro, S., & Djorgovski, S. G. 2003, *ApJ*, 595, L9
- Raftery, A. E. & Lewis, S. M. 1992, in *Bayesian Statistics*, ed. J. M. Bernardo (OUP), 765
- Rigby, J. R., Wuyts, E., Gladders, M., Sharon, K., & Becker, G. 2011, in *American Astronomical Society Meeting Abstracts*, Vol. 217, American Astronomical Society Meeting Abstracts, 114.04
- Roberts, G., Gelman, A., & Gilks, W. 1997, *Annals of Applied Probability*, 110, <http://www.jstor.org/view/10505164/di984007/98p0121o/0>
- Sajina, A., Scott, D., Dennefeld, M., Dole, H., Lacy, M., & Lagache, G. 2006, *MNRAS*, 369, 939
- Schaerer, D. & de Barros, S. 2009, *A&A*, 502, 423
- Schaerer, D. & Vacca, W. D. 1998, *ApJ*, 497, 618
- Serra, P., Amblard, A., Temi, P., Burgarella, D., Giovannoli, E., Buat, V., Noll, S., & Im, S. 2011, *ArXiv e-prints*
- Silva, L., Granato, G. L., Bressan, A., & Danese, L. 1998, *ApJ*, 509, 103
- Storey, P. J. & Hummer, D. G. 1995, *MNRAS*, 272, 41
- Tojeiro, R., Heavens, A. F., Jimenez, R., & Panther, B. 2007, *MNRAS*, 381, 1252
- Vázquez, G. A. & Leitherer, C. 2005, *ApJ*, 621, 695
- Verde, L., Peiris, H. V., Spergel, D. N., Nolte, M. R., Bennett, C. L., Halpern, M., Hinshaw, G., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Page, L., Tucker, G. S., Wollack, E., & Wright, E. L. 2003, *ApJS*, 148, 195
- Walcher, J., Groves, B., Budavári, T., & Dale, D. 2010, *Ap&SS*, 257

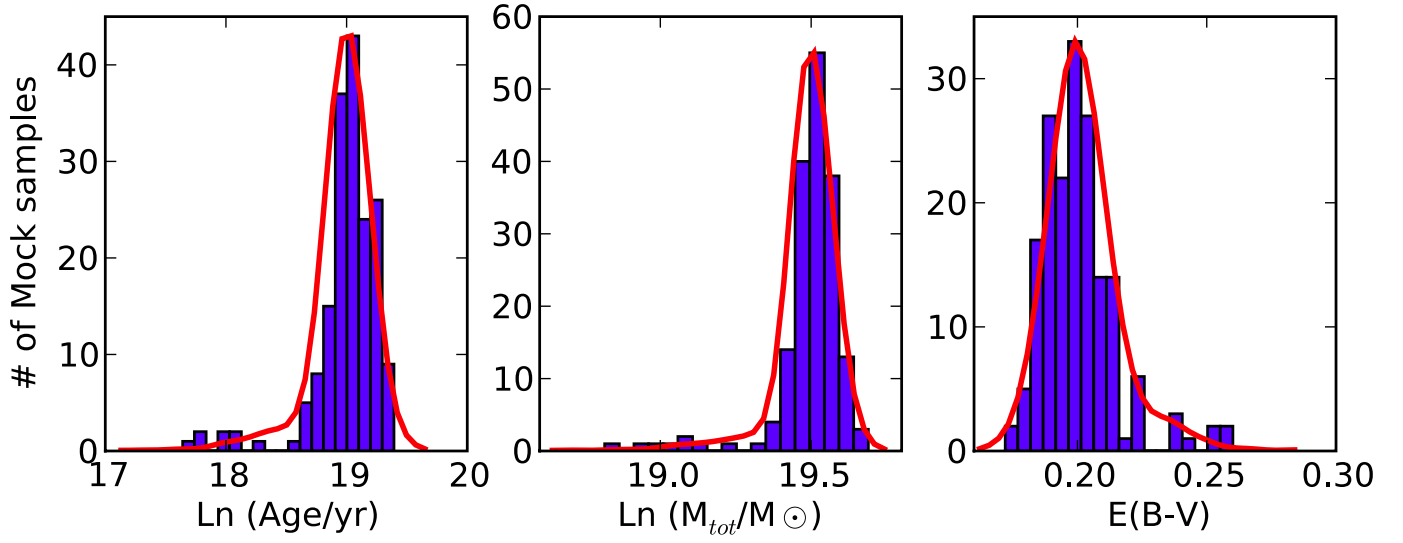


FIG. 1.— The red line shows the posterior probability distribution for the reference model described in the text. The blue histogram represents the distribution of best-fit values obtained from the analysis of 172 Mock data realizations, obtained convolving the reference fluxes with a random Gaussian scatter of amplitude equal to the photometric error of the reference model. The agreement confirms the reliability of our error estimation procedure.

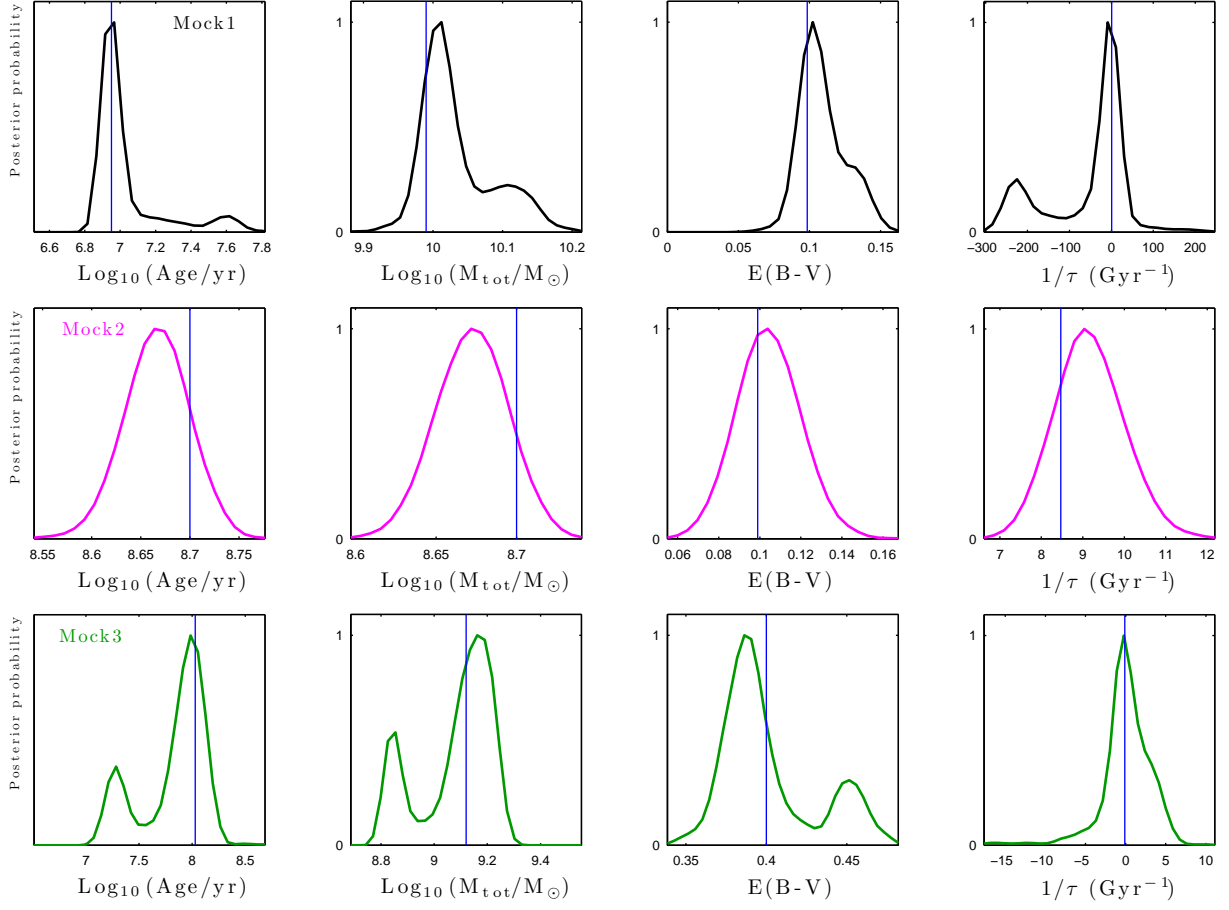


FIG. 2.— Each row shows the marginalized 1-D posterior probability distribution for the SED fitting parameters for one of three illustrative mock catalogs of galaxies with Exponential Star Formation Rate (increasing and decreasing). As discussed in the text, multi-modality in the posterior is often observed for such models. The input parameters are shown as blue vertical lines; in each case the agreement between input and recovered parameters is good.

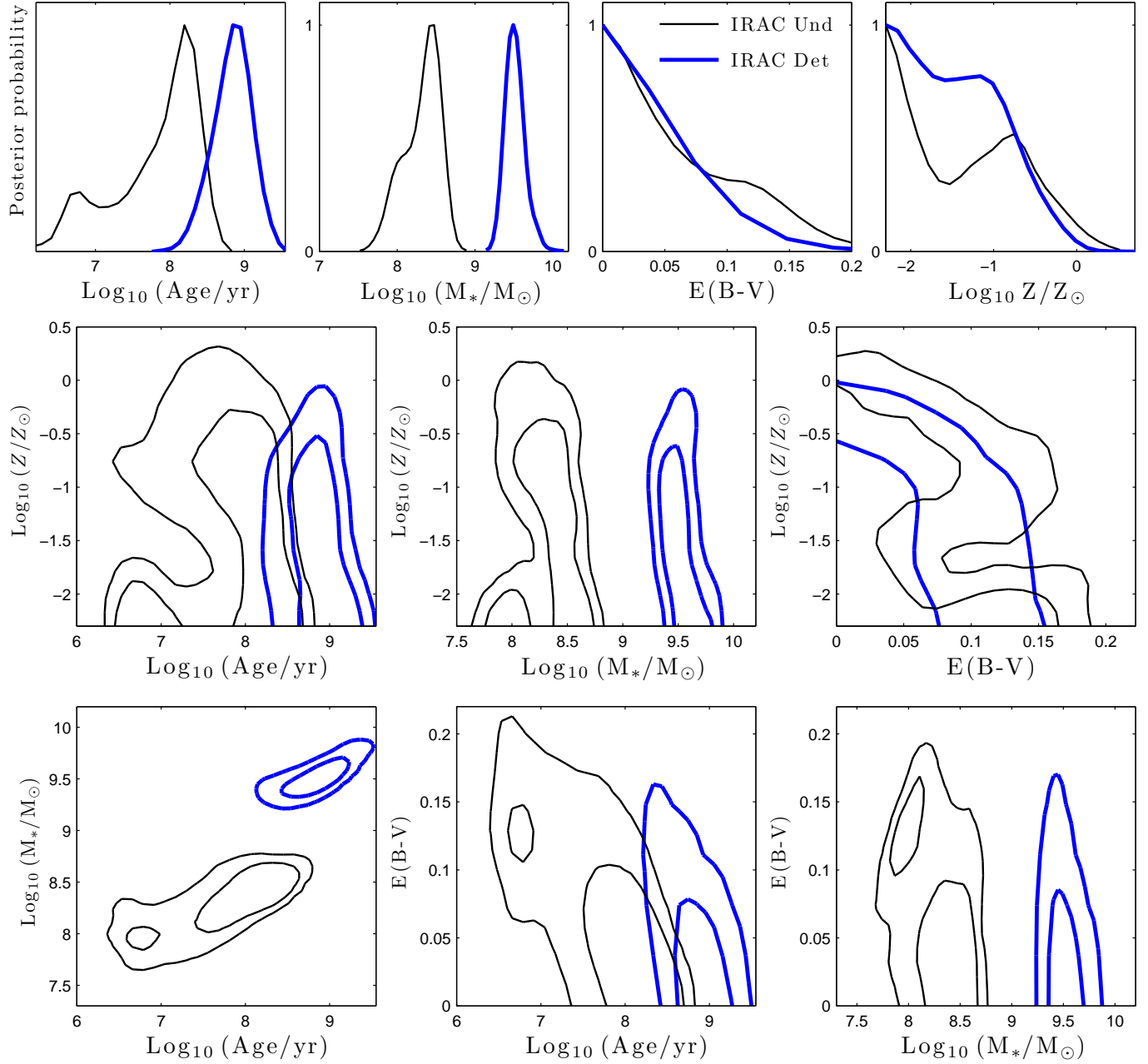


FIG. 3.— **Top row:** 1D marginalized posterior distributions for constant star formation history models, for the IRAC-undetected (black, thin) and IRAC-detected (blue, thick) stacked samples. **Bottom rows:** 2D marginalized contours defining the 68 and 95% confidence regions on the corresponding pairs of parameters.

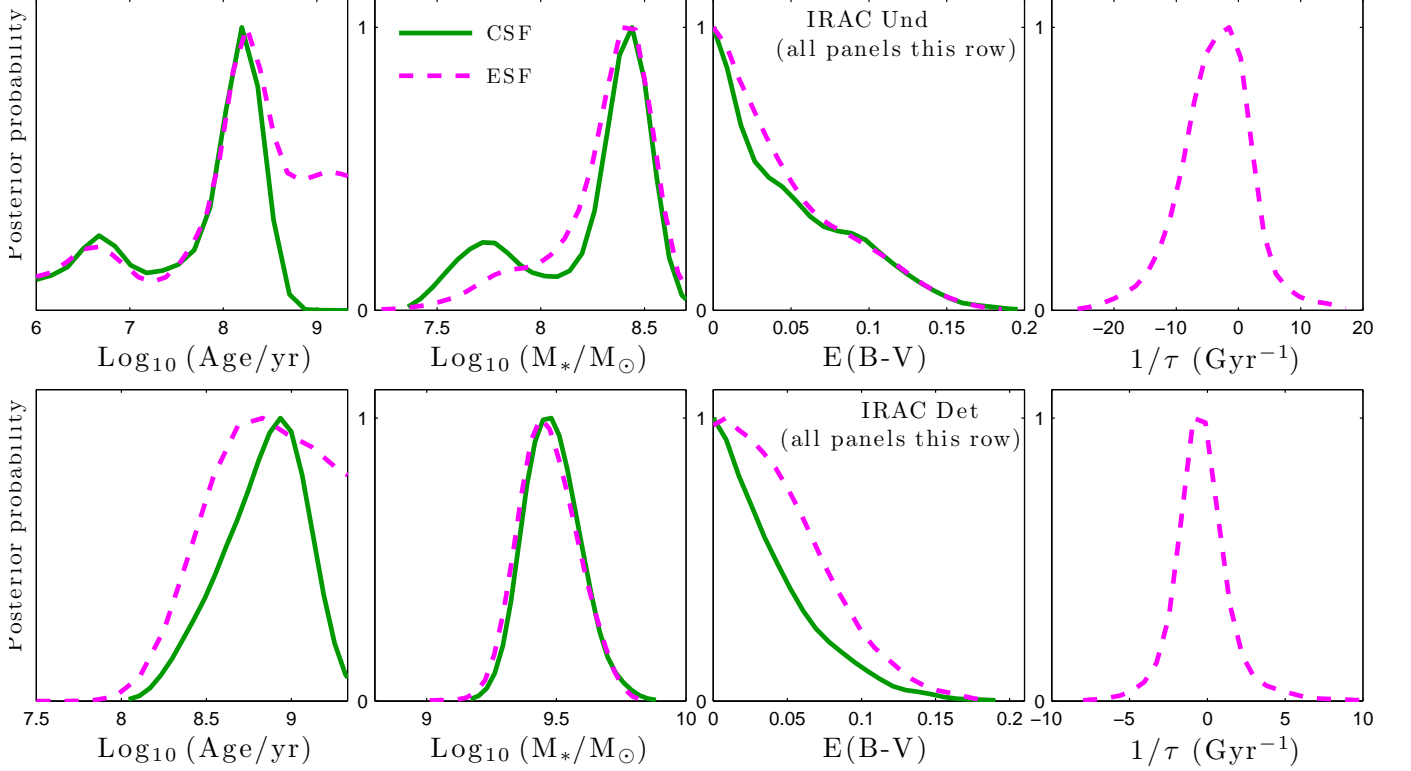


FIG. 4.— **Top row:** Impact of the assumption on the star formation history (constant, green, solid; exponential, magenta, dashed) on the posterior probability distribution of SED fit parameters. For this analysis we held the metallicity fixed at the value  $Z = 0.02 Z_{\odot}$ . The main appreciable difference is the inability to rule out very old ages (comparable to the age of the Universe at  $z = 3.1$ ) for exponential SFH, for both the IRAC Undetected (*top*) and IRAC Detected (*bottom*) samples. Our constraints on  $1/\tau$  show that for both samples the CSF value  $1/\tau = 0$  is well within the region allowed at 68% confidence.

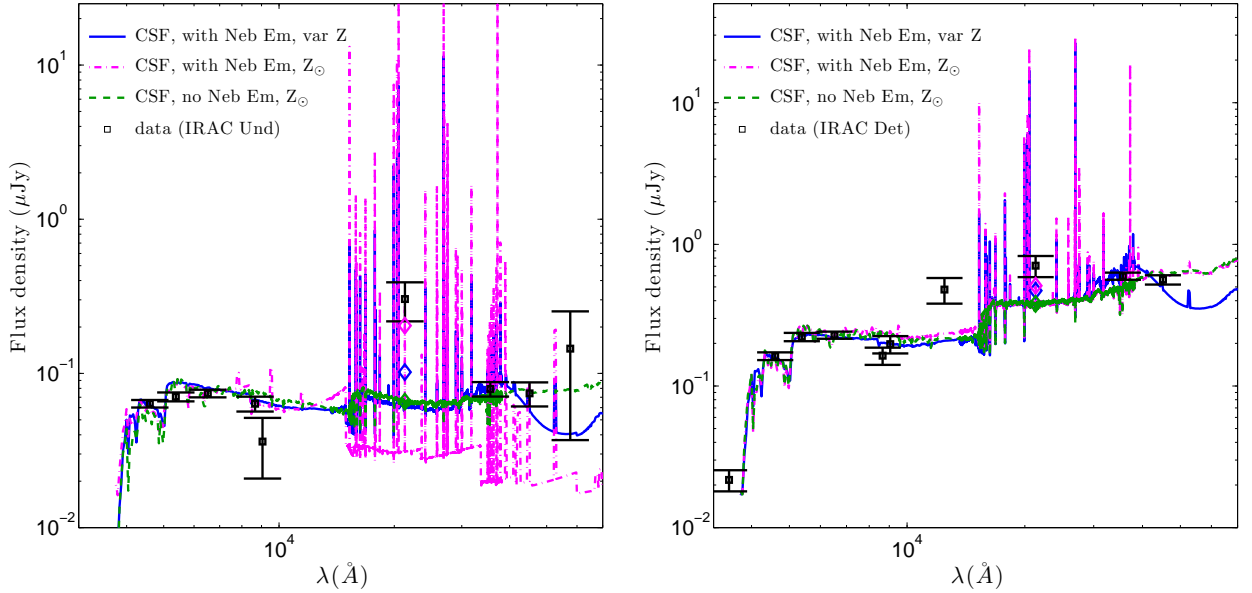


FIG. 5.— Observed-frame data points (black) and best-fit SEDs for constant star formation history and different modeling assumptions presented in Table 3. The colored diamonds show the best fit point for CSF in the K-band, the most sensitive to nebular emission lines. **Left:** IRAC Undetected sample. **Right:** IRAC Detected sample. Data points with negative median stacked fluxes are not shown. The blue solid line shows the best-fit SED from the run with varying metallicity, corresponding to a value of  $Z/Z_{\odot} = 0.005$  for the Detected sample and  $Z/Z_{\odot} = 0.006$  for the Undetected sample.

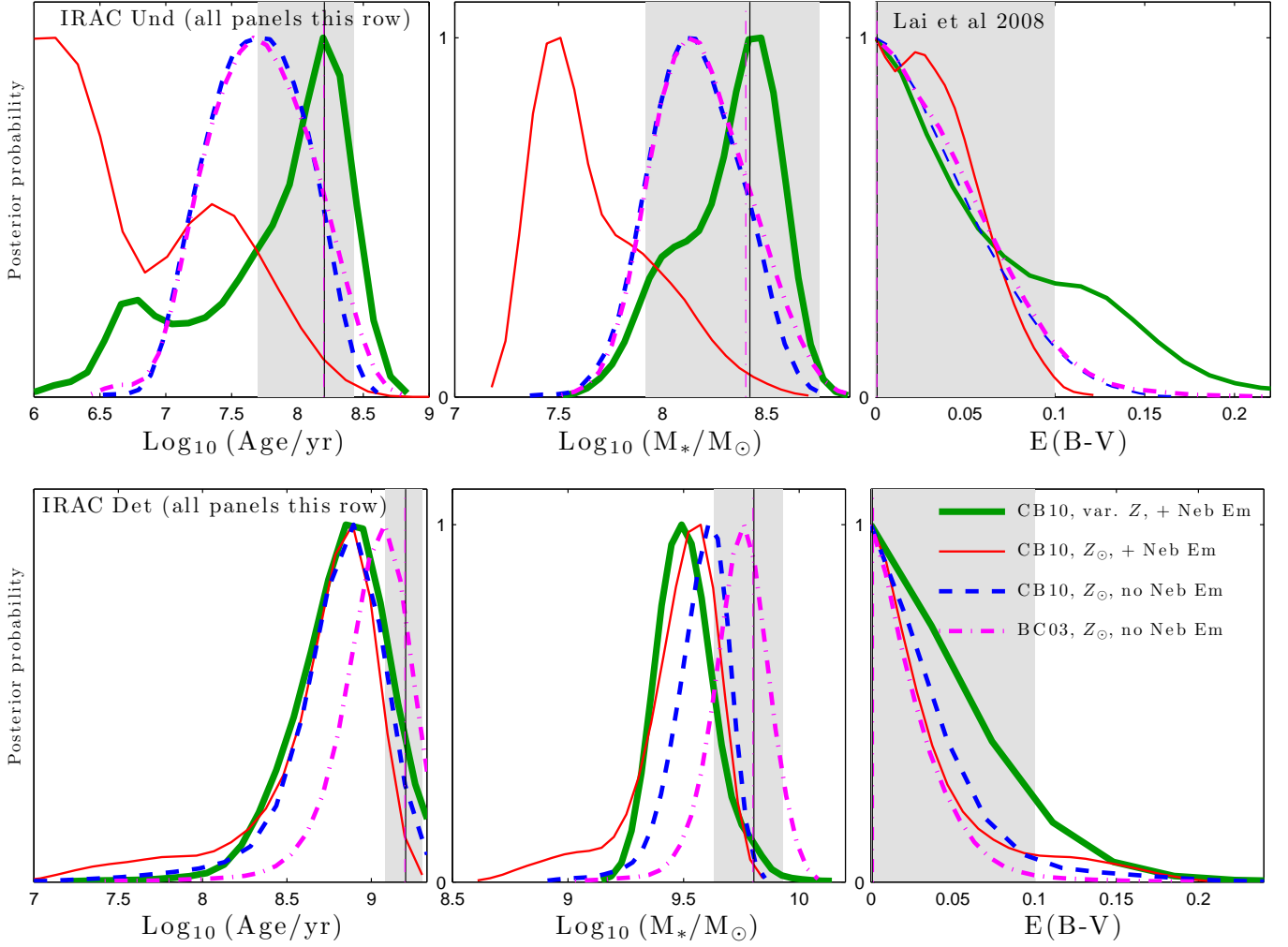


FIG. 6.— Impact of inclusion of nebular emission, metallicity, evolution in SPS templates, and best-fit vs mean values on the posterior probability distribution of SED fit parameters for the IRAC Undetected (*top*) and IRAC Detected (*bottom*) samples. Constant Star Formation History is assumed. The thin black vertical line and the grey shaded area indicate the best-fit models and 68% confidence levels from Lai et al 2008. The magenta dot-dashed vertical lines and curves show the best-fit models and probability distributions of our SED fit for the same SPS models and metallicity. This comparison shows the slight shift resulting from the use of mean values of the posterior distribution as opposed to best fit values. For the Undetected stack, the introduction of nebular emission causes the appearance of a second mode at very young ages and very low masses, which is particularly relevant at solar metallicity, where the contribution of the emission lines is stronger. Results are summarized in Table 3.



Parameter	Symbol	Type	Prior	Range
Age since star formation begun	Age	Input	flat in $\log(\text{Age})$	$\log_{10}(\text{Age/yr}) \in [6.0-9.33]$
Total mass processed into stars	$M_{\text{tot}}$	Input	flat in $\log(M_{\text{tot}})$	$\log_{10}(M_{\text{tot}}/M_{\odot}) \in [4.0-15.0]$
Dust reddening	E(B-V)	Input	flat in E(B-V)	$E(B-V) \in [0.0-1.0]$
E-folding time	$\tau$	Input	flat in $\log(\tau)$	$\tau \in [-4 \text{ Gyr}, 4 \text{ Gyr}]$
Metallicity	$Z/Z_{\odot}$	Input	flat in $\log(Z/Z_{\odot})$	$\log_{10}(Z/Z_{\odot}) \in [-2.3, 0.7]$
Stellar Mass at $t = \text{Age}$	$M_{\star}$	Derived		

TABLE 1

INPUT PARAMETERS OF THE MCMC CODE, PRIORS USED IN THE SED FIT ANALYSIS, AND DERIVED PARAMETERS USED TO REPORT RESULTS. THE E-FOLDING TIME  $\tau$  IS ALWAYS THE INPUT VARIABLE OF THE SPS CODE, BUT WE HAVE CHOSEN USE  $1/\tau$  AS THE MCMC SAMPLING VARIABLE, AS DISCUSSED IN THE TEXT.

Sample	SFH	$Z/Z_{\odot}$	Age (Gyr)	E(B-V)	$M^*$ ( $10^8 M/M_{\odot}$ )	$\tau$ (Gyr)	best fit $\chi^2/\text{d.o.f.}$
IRAC Det	CSF	0.02	0.67[0.38 – 1.2]	0.038[0 – 0.047]	31[24 – 39]	$\infty$	13.4/9
	ESF	0.02	0.83[0.54 – 2.1]	0.046[0 – 0.059]	29[23 – 38]	$-3.0[ \tau  > 0.67]$	10.8/8
IRAC Und	CSF	0.02	0.05 [0.04 - 0.16]	0.049 [0 - 0.064]	1.7 [0.6- 3.0]	$\infty$	7.6/9
	ESF	0.02	0.22 [0.001 - 2.1]	0.047 [0 - 0.059]	2.0 [1.2 - 3.1]	$-2.7 [ \tau  > 0.12]$	5.2/8

TABLE 2

RESULTS OF SED FITTING FOR CONSTANT AND EXPONENTIAL STAR FORMATION HISTORY. WE REPORT MEAN EXPECTATION VALUES AND 68% CONFIDENCE REGIONS OF PARAMETERS. FOR EXPONENTIAL SFH THE PARAMETER  $\tau$  IS CLOSE TO THE CSF VALUE (DEFINED BY  $|\tau| = 4 \text{ Gyr}$  BECAUSE OF THE PRIORS USED IN ESF MODELS). FOR BOTH SAMPLES THERE IS A MODEST IMPROVEMENT IN THE BEST-FIT  $\chi^2$  WHEN USING EXPONENTIAL SFH, BUT CSF IS INCLUDED IN THE RANGE OF VALUES ALLOWED AT 68% CONFIDENCE, SHOWING THAT THE DATA DO NOT FAVOR ESF SIGNIFICANTLY.

Sample	Neb	SPS	$Z/Z_{\odot}$	Age (Gyr)	E(B-V)	$M^*$ ( $10^8 M/M_{\odot}$ )	best fit $\chi^2/\text{d.o.f.}$
IRAC Det	Y	CB10	0.036[0.005 - 0.07]	0.67[0.37 – 1.18]	0.046[0 – 0.05]	32[25 – 42]	13.2/8
	Y	CB10	1	0.47[0.24 – 0.94]	0.039[0 – 0.041]	29 [21 - 42]	18.6/9
	N	CB10	1	0.72[0.33 – 1.3]	0.028[0 – 0.035]	36[28 – 47]	19.6/9
	N	BC03	1	1.0[0.69 – 1.6]	0.025[0 – 0.03]	55[43 – 71]	15.5/9
L08 Det	N	BC03	1	1.6 [1.2 - 2.0]	[0 - 0.1]	64 [43 - 85]	–
IRAC Und	Y	CB10	0.05 [0.005 - 0.13]	0.06 [0.01 - 0.2]	0.066 [0 - 0.085]	2.0 [1.1- 3.4]	6.84/8
	Y	CB10	1	0.007 [0.001 - 0.018]	0.035 [0 - 0.045]	0.47 [0.26 - 0.98]	9.92/9
	N	CB10	1	0.049 [0.021 - 0.12]	0.04 [0 - 0.056]	1.4 [0.89- 2.4]	11.7/9
	N	BC03	1	0.052 [0.02 - 0.13]	0.039 [0 - 0.049]	1.5 [0.92 - 2.7]	11.8/9
L08 Und	N	BC03	1	0.16 [0.05 - 0.27]	0 [0 - 0.1]	2.6 [0.83 - 5.7]	–

TABLE 3

MEAN EXPECTATION VALUES AND 68% CONFIDENCE REGIONS FROM SED FITTING FOR CONSTANT STAR FORMATION HISTORY, USING DIFFERENT ASSUMPTIONS ON INCLUSION OF NEBULAR EMISSION, SPS MODELING AND METALLICITY. SOLAR METALLICITY IS EXCLUDED AT 95% CONFIDENCE BY THE DATA; SEE TEXT FOR MORE DETAILS. FOR COMPARISON, WE ALSO REPORT BEST-FIT VALUES AND 68% CONFIDENCE REGIONS OBTAINED FOR THE SAME SAMPLES BY L08.